ABSTRACT

A method for characterizing a musical recording as a set of scalar descriptors, each of which is based on human perception. A group of people listens to a large number of musical recordings and assigns to each one many scalar values, each value describing a characteristic of the music as judged by the human listeners. Typical scalar values include energy level, happiness, danceability, melodicness, tempo, and anger. Each of the pieces of music judged by the listeners is then computationally processed to extract a large number of parameters which characterize the electronic signal within the recording. Algorithms are empirically generated which correlate the extracted parameters with the judgments based on human perception to build a model for each of the scalars of human perception. These models can then be applied to other music which has not been judged by the group of listeners to give to each piece of music a set of scalar values based on human perception. The set of scalar values can be used to find other pieces that sound similar to humans or vary in a dimension of one of the scalars.